Lecture #2: Full Belief

Kenny Easwaran [1] Branden Fitelson [2]

[1]Philosophy @ USC
easwaran@usc.edu

[2]Philosophy & RuCCS @ Rutgers
&
MCMP @ LMU
branden@fitelson.org

- Review from last time. We're after *formal*, *synchronic*, *epistemic*, *coherence requirements* (of ideal rationality).
- Our strategy is to apply a Joyce-style accuracy-dominance approach to three different types of judgment sets.
- Last time, we examined Joycean arguments for *probabilism* as a CR for *credences*. We prefer to think of these as arguments that *both* $b(\cdot)$ *and* $s(\cdot, E)$ are probabilistic.
  1. The overall *degree* of inaccuracy of $b(\cdot)$ — at w — is (and should be) measured by some function $\mathcal{I}_b(b(\cdot), v_w(\cdot))$, which is *proper* (or, at least, *probabilistically admissible*).
  2. If (1) is true, then $b(\cdot)$ is non-(weakly)-dominated in $\mathcal{I}_b$-accuracy *iff* $b(\cdot)$ is *probabilistic*.
  3. If $b(\cdot)$ is (weakly) dominated in $\mathcal{I}_b$-accuracy, then — *no matter what the total evidence E is* — $b(\cdot) \neq s(\cdot, E)$.
  4. $\therefore$ if (1) is true, then $b(\cdot) = s(\cdot, E)$ — for *some* body of total evidence E — *only if* $b(\cdot)$ — and $\therefore$ $s(\cdot, E)$ — is probabilistic.

- Premise (1) seems to (inevitably) smuggle-in some probabilistic assumptions (*via propriety/*Pr*-admissibility*). In this sense, such arguments seem "circular" (to some extent).
  - ☞ From now on, I will simply *assume* that $s(\cdot, E)$ is always a *probability* function, *i.e.*, that (PRE) is correct.
- However, the argument above is still interesting because it introduces the following key theoretical notion.

  Evidential Consistency (EC). A judgment set is *evidentially consistent* just in case there exists *some* body of total evidence E which supports each of its members.
- (EC) allows us to restate (3) & (4) in the above argument as:
  3. If $b(\cdot)$ is (weakly) dominated in $\mathcal{I}_b$-accuracy, then $b(\cdot)$ is *evidentially inconsistent*.
  4. $\therefore$ if (1) is true, then $b(\cdot)$ is evidentially consistent *only if* $b(\cdot)$ — and $\therefore$ $s(\cdot, E)$ — is probabilistic.
- (EC) will play a key role in the application to full belief.

- Here is a — perhaps *the* — "paradigm" CR [28, 30, 27, 22].
  - **The Consistency Requirement for Belief Sets** (CB). All belief *sets* should be *logically (viz., alethically) consistent*.
- (CB) follows from the following (narrow-scope) *norm*:
  - **The Truth Norm for Belief** (TB). All (individual) beliefs should (alethically–ideally) be *true*.
- Alethic norms [(CB)/(TB)] can conflict with evidential norms.
  - **The Evidential Norm for Belief** (EB). All (individual) beliefs should (evidentially–ideally) be *supported by the evidence*.
- In some cases (*e.g.*, preface cases), agents satisfy (EB) while violating (CB) — this generates an alethic/evidential *conflict*.
- Such alethic/evidential conflicts needn't give rise to states that receive an (overall) evaluation as *ir*rational (nor must they inevitably give rise to rational *dilemmas*) [6, 24, 14, 23].
- We'll refer to the claim that there exist *some* such cases as *the datum*. Foley's [14] explanation of *the datum* is helpful.

"…if the avoidance of recognizable inconsistency were an absolute prerequisite of rational belief, we could not rationally believe each member of a set of propositions and also rationally believe of this set that at least one of its members is false. But this in turn pressures us to be unduly cautious. It pressures us to believe only those propositions that are certain or at least close to certain for us, since otherwise we are likely to have reasons to believe that at least one of these propositions is false. At first glance, the requirement that we avoid recognizable inconsistency seems little enough to ask in the name of rationality. It asks only that we avoid certain error. It turns out, however, that this is far too much to ask."

- We will not argue for *the datum* here. We think Foley [14], Christensen [6], Kolodny [24], and others have made a compelling case for it. Today, it is our *point of departure*. [But, we do have a new Preface case that we think is definitive.]

- Here is a Preface case that need not involve higher-order evidence (*via* reflection on one's own fallibility).

Homogeneous Preface Paradox. John is an excellent empirical scientist. He has devoted his entire (long and esteemed) scientific career to gathering and assessing the evidence that is relevant to the following first-order, empirical hypothesis: (H) all scientific/empirical books of sufficient complexity contain at least one false claim. By the end of his career, John is ready to publish his masterpiece, which is an exhaustive, encyclopedic, 15-volume (scientific/empirical) book which aims to summarize (all) the evidence that contemporary empirical science takes to be relevant to H. John sits down to write the Preface to his masterpiece. Rather than reflecting on his own fallibility, John simply reflects on the contents of (the main text of) his book, which constitutes *very strong inductive evidence in favor of* H. On this basis, John (inductively) infers H. But, John also believes each of the individual claims asserted in the main text of the book. Thus, because John believes (indeed, knows) that his masterpiece instantiates the antecedent of H, the (total) set of John's (rational/justified) beliefs is inconsistent.

☞ While John's belief set (on the preface agenda) is *alethically* inconsistent, it is *not evidentially* inconsistent.

- Some philosophers construe *the datum* as reason to believe that ($\star$) *there are no coherence requirements for full belief*.
- Christensen [6] thinks (a) *credences* do have coherence requirements (*probabilism*); ($\star$) full beliefs do *not*; (b) what *seem* to be CRs for full belief can be explained *via* (a).
- Kolodny [24] agrees with ($\star$), but he disagrees with (a) and (b). He thinks (c) full belief is *explanatorily indispensable*; (d) there are *no* coherence requirements for *any* judgments; (e) what *seem* to be CRs for full belief can be explained *via* (EB).
- Christensen & Kolodny *agree* — *trivially*, *via* ($\star$) — that:

  (†) *If* there are *any* coherence requirements for full belief, *then* (CB) is a coherence requirement for full belief.

- We [2, 11] agree with Christensen on (a) and Kolodny on (c), but we disagree with them on ($\star$), (d), (e), and (†). We'll explain how to ground "conflict-proof" CRs for full belief, by analogy with Joyce's [21, 19] argument(s) for probabilism.

- We begin with some background assumptions/notation.
  - $B(p) \overset{\text{def}}{=} S$ believes that p. $D(p) \overset{\text{def}}{=} S$ disbelieves that p.
- S makes judgments regarding propositions in a (finite) *agenda* ($\mathcal{A}$) of (classical, possible-worlds) propositions. We'll use "**B**" to denote the *set* of S's judgments on $\mathcal{A}$.

☞ We're only evaluating *explicit judgments* (*on* $\mathcal{A}$) — we assume nothing about off-$\mathcal{A}$ ("implicit") commitments.

- We'll assume the following about B/D on $\mathcal{A}$. The first assumption is integral to the framework. The second two assumptions are made for simplicity (and could be relaxed).
  - **Accuracy conditions**. $B(p)$ [$D(p)$] is accurate iff p is T [F].

  - **Incompatibility**. $B(p) \Rightarrow \neg D(p)$.
  - **Opinionation**. $B(p) \vee D(p)$.
- See Extras (26) for Kenny's [10] relaxation of Opinionation.
- We assume *belief/world independence* — until slide (17).

- Now, we can explain how our new CRs were discovered, by analogy with Joyce's [21, 19] argument(s) for probabilism.
- Both arguments can be seen as involving three key steps.
- **Step 1**: Define $\mathring{\mathbf{B}}_w$ — the *vindicated (viz., alethically ideal* or *perfectly accurate)* judgment set (on $\mathcal{A}$), at world w.
  - $\mathring{\mathbf{B}}_w$ contains $B(p)$ [$D(p)$] iff p is true (false) at w.
    - Heuristically, we can think of $\mathring{\mathbf{B}}_w$ as the set of judgments that an omniscient agent would have (on $\mathcal{A}$, at w).
- **Step 2**: Define $d(\mathbf{B}, \mathring{\mathbf{B}}_w)$ — a measure of distance between $\mathbf{B}$ and $\mathring{\mathbf{B}}_w$. That is, a measure of $\mathbf{B}$'s *overall inaccuracy* (at w).
  - $d(\mathbf{B}, \mathring{\mathbf{B}}_w) \overset{\text{def}}{=}$ the number of inaccurate judgments in $\mathbf{B}$ at w.
    - *Hamming distance* [8] between the binary vectors $\mathbf{B}$, $\mathring{\mathbf{B}}_w$.
- **Step 3**: Adopt a *fundamental epistemic principle*, which uses $d(\mathbf{B}, \mathring{\mathbf{B}}_w)$ to ground a coherence requirement for $\mathbf{B}$.
- This last step is the philosophically crucial one...

- Given our choices at Steps 1 and 2, there is ***a*** choice we can make at Step 3 that will yield (CB) as a requirement for $\mathbf{B}$.

  Possible Vindication (PV). There exists some possible world w at which all of the judgments in $\mathbf{B}$ are accurate. Or, to put this more formally, in terms of d: $(\exists w)[d(\mathbf{B}, \mathring{\mathbf{B}}_w) = 0]$.

- Possible vindication is *one* way we could go here. But, our framework is much more general than the classical one. It allows for many other choices of fundamental principle.
- Like Joyce [21, 19] — who makes the analogous move with credences, to ground probabilism — we retreat from (PV) to the weaker: *avoidance of (weak) dominance in* $d(\mathbf{B}, \mathring{\mathbf{B}}_w)$.

  Weak Accuracy-Dominance Avoidance (WADA).

  There does *not* exist an alternative belief set $\mathbf{B}'$ such that:
  (i) $(\forall w)[d(\mathbf{B}', \mathring{\mathbf{B}}_w) \leq d(\mathbf{B}, \mathring{\mathbf{B}}_w)]$, and
  (ii) $(\exists w)[d(\mathbf{B}', \mathring{\mathbf{B}}_w) < d(\mathbf{B}, \mathring{\mathbf{B}}_w)]$.

- Completing Step 3 in this way reveals new CRs for $\mathbf{B}$...

- Ideally, we want a coherence requirement that [like (CB)] can be motivated by considerations of *accuracy* (*viz.*, a CR that is *entailed by* alethic requirements such as TB/CB/PV).
- ☞ But, in light of (*e.g.*) preface cases, we also want a CR that is *weaker* than (CB). More precisely, we want a CR that is weaker than (CB) *in such a way that it is also entailed by* (EB).
- We can show that our new CRs [*e.g.*, (WADA)] fit the bill, *if* we assume the following "probabilistic-evidentialist" necessary condition for the satisfaction of (EB).

  Necessary Condition for Satisfying (EB). $\mathbf{B}$ satisfies (EB), *i.e.*, all judgments in $\mathbf{B}$ are *supported by the evidence*, ***only if***:

  ($\mathcal{R}$) There exists *some* Pr-function that probabilifies (*i.e.*, assigns Pr greater than $1/2$ to) each belief in $\mathbf{B}$ and dis-probabilifies (*i.e.*, assigns Pr less than $1/2$ to) each disbelief in $\mathbf{B}$.

- "Probabilistic-evidentialists" will disagree about *which* $\Pr(\cdot)$ undergirds (EB) [5, 32, 15, 20]; but, they agree on (EB) $\Rightarrow$ ($\mathcal{R}$). In fact, Pr-evidentialists will agree that (EB) $\Rightarrow$ (EC) $\Rightarrow$ ($\mathcal{R}$).

- Here are the logical relationships between key norms:

Truth Norm for Belief:	(TB)
	$\Downarrow\ \not\Uparrow$
Consistency Norm for Belief (*viz.*, PV):	(CB)/(PV)
	$\Downarrow\ \not\Uparrow$
☞ Weak Accuracy-Dominance Avoidance:	$\boxed{\text{(WADA)}}$
	$\Uparrow\ \not\Downarrow$
Evidential Norm for Belief:	(EB)

$$
\begin{tikzcd}
\text{(TB)} \arrow[d, Rightarrow] & & \text{(EB)} \arrow[d, Rightarrow] \\
\text{(CB)/(PV)} \arrow[dr, Rightarrow] & & \text{(EC)} \arrow[dl, Rightarrow] \\
& (\mathcal{R}) \arrow[d, Rightarrow] & \\
& \text{(WADA)} &
\end{tikzcd}
$$

- See slide #20 for a bigger map w/11 requirements/norms.

- There are many advantages to adopting ($\mathcal{R}$), rather than (WADA), as our (ultimate) CR for full belief. Here are a few:
  - First, (WADA) is (intuitively) *too weak* to serve as our (ultimate) CR — $\{B(p), B(\neg p)\}$ may be *non-dominated*, as the following table reveals (*ditto* for $\{D(p), D(\neg p)\}$).

	P	$\neg P$	$B(P)$	$B(\neg P)$	$B(P)$	$D(\neg P)$	$D(P)$	$B(\neg P)$	$D(P)$	$D(\neg P)$
w_1	**F**	**T**	−	+	−	−	+	+	+	−
w_2	**T**	**F**	+	−	+	+	−	−	−	+

  - ($\mathcal{R}$) ⇒ (NCP) $D(p) \equiv B(\neg p)$, which *rules-out* $\{B(p), B(\neg p)\}$.
  - ($\mathcal{R}$) is strictly stronger than (WADA) + (NCP). Indeed, we conjecture that ($\mathcal{R}$) is the strongest CR (uncontroversially) entailed by both alethic and evidential considerations.
  - (WADA) only makes sense for *finite* agendas, whereas ($\mathcal{R}$) is potentially applicable to *infinite* agendas (if there be such).
  - ($\mathcal{R}$) entails (WADA$_d$), for *any additive inaccuracy measure d*. In this sense, ($\mathcal{R}$) is *robust* across choices of d.

- Proof of the central result that ($\mathcal{R}$) ⇒ (WADA).

Let Pr be a probability function that represents **B** in sense of ($\mathcal{R}$). Consider the *expected* inaccuracy of a belief set — the sum of $\Pr(w) \cdot d(\mathbf{B}, \mathring{\mathbf{B}}_w)$. Since $d(\mathbf{B}, \mathring{\mathbf{B}}_w)$ is a sum of components for each proposition (1 if **B** disagrees with w on the proposition and 0 if they agree), and since expectations are linear, the expected inaccuracy is the sum of the expectation of these components. The expectation of the component for disbelieving p is $\Pr(p)$ while the expectation of the component for believing p is $1 - \Pr(p)$. Thus, if $\Pr(p) > 1/2$ then believing p is the attitude that uniquely minimizes the expectation, while if $\Pr(p) < 1/2$ then disbelieving p is the attitude that uniquely minimizes the expectation. Thus, since Pr represents **B**, this means that **B** has strictly lower expected inaccuracy than any other belief set with respect to Pr. Suppose, for *reductio*, that some **B**′ (weakly) dominates **B**. Then, **B**′ must be no farther from vindication than **B** in any world, and thus **B**′ must have expected inaccuracy no greater than that of **B**. But **B** has strictly lower expected inaccuracy than any other belief set. Contradiction. ∴ **B** must be non-dominated.

- The key to our central theorem that ($\mathcal{R}$) ⇒ (WADA) is that our inaccuracy measure $d(\mathbf{B}, \mathring{\mathbf{B}}_w)$ is *evidentially proper*.

Definition (Evidential Propriety)

Suppose a judgment set **J** of type $\mathfrak{J}$ is supported by the evidence. That is, suppose there exists some evidential probability function $\Pr(\cdot)$ which represents **J** (in the appropriate sense of "represents" for sets of type $\mathfrak{J}$). If this is sufficient to ensure that **J** minimizes expected inaccuracy (relative to Pr), according to the measure of inaccuracy $\mathfrak{I}(\mathbf{J}, \mathring{\mathbf{J}}_w)$, then we will say that the measure $\mathfrak{I}$ is **evidentially proper**.

- If an inaccuracy measure is evidentially *im*proper, then some probabilistically representable judgment sets will be *ruled out* as *irrational via* accuracy-dominance (WADA).
- This would engender a *conflict* between alethic and evidential requirements for judgment, which is exactly what coherence requirements are *not* supposed to do.
- In our book [12], evidential propriety plays a central role.

- It is useful to draw an analogy between the norms and requirements we've been discussing, and principles in rational choice theory. The Decision-Theoretic Analogy.

Epistemic Principle	Analogous Decision-Theoretic Principle
(TB)	(AMU) Do ϕ only if ϕ maximizes utility in the *actual* world.
(CB)	(PMU) Do ϕ only if ϕ maximizes u in *some possible* world.
($\mathcal{R}$)	(MEU) Do ϕ only if ϕ maximizes EU (relative to *some* Pr).
(WADA)	(WDOM) Do ϕ only if ϕ is *not weakly dominated* in utility.
(SADA)	(SDOM) Do ϕ only if ϕ is *not strictly dominated* in utility.

- Like (TB), (AMU) is *not* a *requirement of rationality*; and, like (CB), (PMU) isn't a rational requirement either. Moreover, also like (CB), seeing this requires "paradoxical" cases [25].
- As Foley (*op. cit.*) explains, (CB) is *too demanding*. But, ($\mathcal{R}$) and (WADA) are *not* — they do *not* "pressure us to believe only those propositions that are (close to) certain for us".

- Michael Caie [4] writes about an example involving self-reference, which causes problems for Joyce-style (accuracy-dominance) arguments for *probabilism*.
- There are analogous examples for full belief. Consider:

  (P) S does not believe that P. [$\neg B(\ulcorner P \urcorner)$.]

- One can argue (Caie-style) that the only non-dominated (opinionated) belief sets on $\{P, \neg P\}$ are $\{B(P), B(\neg P)\}$ and $\{D(P), D(\neg P)\}$, which are both *ruled-out* by ($\mathcal{R}$).

	P	$\neg P$	$B(P)$	$B(\neg P)$	$B(P)$	$D(\neg P)$	$D(P)$	$B(\neg P)$	$D(P)$	$D(\neg P)$
w_1	**F**	**T**	−	+	−	−	×	×	×	×
w_2	**T**	**F**	×	×	×	×	−	−	−	+

- The "×"s indicate that these worlds are *ruled-out* (*a priori*) by the definition of P. As such, the only non-dominated belief sets seem to be $\{B(P), B(\neg P)\}$ and $\{D(P), D(\neg P)\}$.
- If this Caie-style reasoning is correct, then it shows that *some of our assumptions must go*. But, which one(s)?

$\mathcal{B}$	$\mathbf{B}_1$	$\mathbf{B}_2$
$\neg X \,\&\, \neg Y$	D	D
$X \,\&\, \neg Y$	D	D
$X \,\&\, Y$	D	D
$\neg X \,\&\, Y$	D	D
$\neg Y$	D	D
$X \equiv Y$	D	D
$\neg X$	B	B
X	B	D
$\neg(X \equiv Y)$	D	D
Y	D	D
$X \vee \neg Y$	B	B
$\neg X \vee \neg Y$	B	B
$\neg X \vee Y$	B	B
$X \vee Y$	D	B
$X \vee \neg X$	B	B
$X \,\&\, \neg X$	D	D

We have the following four facts regarding $\mathbf{B}_1$ & $\mathbf{B}_2$:

(1) $\mathbf{B}_1$ is *weakly* dominated in inaccuracy by $\mathbf{B}_2$ (this is easily verified by simple counting). Thus, $\mathbf{B}_1$ *violates* (WADA).

(2) $\mathbf{B}_1$ is *not strictly* dominated in inaccuracy by *any* belief set over $\mathcal{B}$ (this can be verified *via exhaustive search* on the set of all belief sets over $\mathcal{B}$). Thus, $\mathbf{B}_1$ *satisfies* (SADA).

(3) $\mathbf{B}_2$ is *not weakly* dominated (in inaccuracy) by *any* belief set over $\mathcal{B}$ (this can be verified *via exhaustive search* on the set of all belief sets over $\mathcal{B}$). Thus, $\mathbf{B}_2$ *satisfies* (WADA).

(4) $\mathbf{B}_2$ is *not* represented (in the sense of Definition 2) by *any* probability function on $\mathcal{B}$, since the set $\mathbf{B}_2$ contains two contradictory pairs: $\{D(Y), D(\neg Y)\}$, $\{D(X \equiv Y), D(\neg(X \equiv Y))\}$. Therefore, $\mathbf{B}_2$ *violates* ($\mathcal{R}$).

(TB) S ought believe p just in case p is true.

(PV) $(\exists w)[d(\mathbf{B}, \mathring{\mathbf{B}}_w) = 0]$. (**B** is *logically/alethically consistent.*)

(SADA) $\nexists\, \mathbf{B}'$ such that: $(\forall w)[d(\mathbf{B}', \mathring{\mathbf{B}}_w) < d(\mathbf{B}, \mathring{\mathbf{B}}_w)]$.

($\mathbf{NW_2S}$) $\nexists\, \beta \subseteq \mathbf{B}$ s.t.: $(\forall w)$ [$>$ 1/2 *of the members of* β *are inaccurate at* w].

($\boldsymbol{\mathcal{R}_r}$) $\exists$ a probability function $\Pr(\cdot)$ such that, $\forall p \in \mathcal{A}$:

$$B(p) \text{ iff } \Pr(p) > r, \text{ and } D(p) \text{ iff } \Pr(p) < 1 - r.$$

(EB) S ought believe p just in case p is supported by S's evidence.

Note: this assumes *only* $(\exists \Pr)(\forall p)\, [\Pr(p) > 1/2 \text{ iff } B(p)]$.

(NWS) $\nexists\, \beta \subseteq \mathbf{B}$ s.t.: $(\forall w)$ [$\geq$ 1/2 *of the members of* β *are inaccurate at* w] & $(\exists w)$ [$>$ 1/2 *of the members of* β *are inaccurate at* w].

(WADA) $\nexists\, \mathbf{B}'$ s.t.: $(\forall w)[d(\mathbf{B}', \mathring{\mathbf{B}}_w) \leq d(\mathbf{B}, \mathring{\mathbf{B}}_w)] \,\&\, (\exists w)[d(\mathbf{B}', \mathring{\mathbf{B}}_w) < d(\mathbf{B}, \mathring{\mathbf{B}}_w)]$.

($\mathbf{NW_1S}$) $\nexists\, \beta \subseteq \mathbf{B}$ s.t.: $(\forall w)$ [$\geq$ 1/2 *of the members of* β *are inaccurate at* w].

(NCP) S disbelieves p iff S believes $\neg p$ [*i.e.*, $D(p) \equiv B(\neg p)$].

- Here is what the logical relations look like, among all of the 11 requirements & norms for (opinionated) full belief.

$$\begin{tikzcd}
 & \text{(TB)} \arrow[d, Rightarrow] & & \text{(EB)} \arrow[d, Rightarrow] \\
 & \text{(CB)/(PV)} \arrow[dr, Rightarrow, "(\mathcal{R}_r)"] & & \text{(EC)} \arrow[dl, Rightarrow] \\
\text{(NW}_2\text{S)} \arrow[ddd, Rightarrow] & & (\mathcal{R})/(\mathcal{R}_{1/2}) \arrow[ll, Rightarrow] \arrow[d, Rightarrow] & \\
 & & \text{(WADA) \& (NCP)} \arrow[dd, Rightarrow] & \\
 & & & \\
\text{(NWS)} \arrow[rr, Leftrightarrow] \arrow[d, Rightarrow] & & \text{(WADA)} \arrow[d, Rightarrow] & \\
\text{(NW}_1\text{S)} \arrow[rr, Leftrightarrow] & & \text{(SADA)} &
\end{tikzcd}$$

- Proof of the claim that (NWS) ⇔ (WADA).

  (⇐) We'll prove the contrapositive. Suppose that some $\mathbf{S} \subseteq \mathbf{B}$ is a witnessing set. Let $\mathbf{B}'$ agree with $\mathbf{B}$ on all judgments outside $\mathbf{S}$ and disagree with $\mathbf{B}$ on all judgments in $\mathbf{S}$. By the definition of a witnessing set, $\mathbf{B}'$ weakly dominates $\mathbf{B}$ in inaccuracy $[d(\mathbf{B}, \mathring{\mathbf{B}}_w)]$.

  (⇒) [Contrapositive again.] Suppose $\mathbf{B}$ is dominated, *i.e.*, that there is some $\mathbf{B}'$ that weakly dominates $\mathbf{B}$ in inaccuracy $[d(\mathbf{B}, \mathring{\mathbf{B}}_w)]$. Let $\mathbf{S} \subseteq \mathbf{B}$ be the set of judgments on which $\mathbf{B}$ and $\mathbf{B}'$ disagree. Then, $\mathbf{S}$ is a witnessing set.

- A similar proof can be given for: ($\mathrm{NW_1S}$) ⇔ (SADA).
- We also know that ($\mathcal{R}$) ⇒ ($\mathrm{NW_2S}$). See next slide for a proof.
- The converse remains an *open question*. That is:

  (O) $(\mathrm{NW_2S}) \overset{?}{\Rightarrow} (\mathcal{R})$.

- We *do* know that ($\mathcal{R}$) is *strictly stronger* than the conjunction (WADA) & (NCP). See the slide #23 for a proof.

Theorem

($\mathcal{R}$) ⇒ ($\mathrm{NW_2S}$).

Proof.

In our proof (slide #14) of the claim that ($\mathcal{R}$) ⇒ (WADA), we established that if Pr represents $\mathbf{B}$, then $\mathbf{B}$ has strictly lower expected inaccuracy than any other belief set with respect to Pr. Assume, for *reductio*, that $\mathbf{S} \subseteq \mathbf{B}$ is a witnessing$_2$ set for $\mathbf{B}$. Let $\mathbf{B}'$ agree with $\mathbf{B}$ on all judgments outside $\mathbf{S}$ and disagree with $\mathbf{B}$ on all judgments in $\mathbf{S}$. Then by the definition of a witnessing$_2$ set, $\mathbf{B}'$ must be no farther from vindication than $\mathbf{B}$ in any world. But this contradicts the fact that $\mathbf{B}$ has strictly lower expected inaccuracy than $\mathbf{B}'$ with respect to Pr. So the witnessing$_2$ set must not exist. □

Theorem

(NDB & NCP) ⇏ ($\mathcal{R}$). [In other words, (WADA & NCP) ⇏ ($\mathcal{R}$).]

Proof.

Let there be six possible worlds, $w_1, w_2, w_3, w_4, w_5, w_6$. And, let $\mathcal{A} \stackrel{\text{def}}{=} \{p_1, p_2, p_3, p_4\}$, where the p_i are defined as follows.

$$p_1 \stackrel{\text{def}}{=} \{w_1, w_2, w_3\} \qquad p_2 \stackrel{\text{def}}{=} \{w_1, w_4, w_5\}$$

$$p_3 \stackrel{\text{def}}{=} \{w_2, w_4, w_6\} \qquad p_4 \stackrel{\text{def}}{=} \{w_3, w_5, w_6\}$$

Let $\mathbf{B} \stackrel{\text{def}}{=} \{B(p_1), B(p_2), B(p_3), B(p_4)\}$. $\mathbf{B}$ is a witnessing$_2$ set, since, in every w_i, *exactly half* of the beliefs in $\mathbf{B}$ are accurate. So, by ($\mathcal{R}$) ⇒ ($\mathrm{NW_2S}$), $\mathbf{B}$ *violates* ($\mathcal{R}$). But, $\mathbf{B}$ *satisfies* (NDB), since every belief set on $\mathcal{A}$ has an expected inaccuracy of 2, relative to the uniform Pr-distribution, which implies that no belief set on $\mathcal{A}$ dominates any other belief set on $\mathcal{A}$. Finally, $\mathbf{B}$ satisfies (NCP), since every pair of beliefs in $\mathbf{B}$ is consistent. □

Parametric Family of Requirements Between ($\mathcal{R}$) and (CB)

($\mathcal{R}_r$) There is a probability function Pr such that, for all $p \in \mathcal{A}$:

  (i) $\mathbf{B}$ contains $B(p)$ iff $\Pr(p) > r$, and

  (ii) $\mathbf{B}$ contains $D(p)$ iff $\Pr(p) < 1 - r$,

  where $r \in [1/2, 1)$.

- Let $\mathbb{B}_n$ denote the class of minimal inconsistent belief sets of size n — each member of $\mathbb{B}_n$ is an inconsistent judgment set of size n containing no inconsistent proper subset.
- Let $\mathbf{B}_n$ be a member of $\mathbb{B}_n$, *i.e.*, $\mathbf{B}_n$ consists of n propositions, there is no world in which all of these n propositions are true, but for each proper subset $\mathbf{B} \subset \mathbf{B}_n$ there is a world in which all members of $\mathbf{B}$ are true.

Theorem

For all $n \geq 2$, if $r \geq \frac{n-1}{n}$ then ($\mathcal{R}_r$) rules out each member of $\mathbb{B}_n$, while if $r < \frac{n-1}{n}$, then ($\mathcal{R}_r$) rules out no member of $\mathbb{B}_n$.

Proof.

Let $\mathbf{B}_n \overset{\text{def}}{=} \phi_1, \ldots, \phi_n$. Let each w_i be a world in which ϕ_i is false, but all other members of $\mathbf{B}_n$ are true. Let Pr be the probability function that assigns $1/n$ to each world w_i and 0 to all other worlds. If $r < n-1/n$, then Pr shows $\mathbf{B}_n$ satisfies $(\mathcal{R}_r)$. This establishes the second half of the Theorem.

For the first half of the Theorem, we proceed *via reductio*. Suppose (for *reductio*) $\mathbf{B}_n$ is a member of $\mathbb{B}_n$ that is *not* ruled out by $(\mathcal{R}_{n-1/n})$. Then there must be some Pr such that for each i, $\Pr(\phi_i) > n-1/n$. Therefore, for each i, $\Pr(\neg\phi_i) < 1/n$. Now, since the disjunction of finitely many propositions is at most as probable as the sum of their individual probabilities, we must have $\Pr(\neg\phi_1 \vee \ldots \vee \neg\phi_n) < 1$. But, since $\mathbf{B}_n$ is inconsistent, $\neg\phi_1 \vee \ldots \vee \neg\phi_n$ is a tautology, and must have probability 1. Contradiction. So $\mathbf{B}_n$ must be ruled out by $(\mathcal{R}_{n-1/n})$. □

- Kenny has written a paper [10] that explains how to relax the assumption of Opinionation in our framework.
- Our approach is equivalent to assigning (in)accurate judgments a *score* of (-1) $+1$, and calculating the *total score* of $\mathbf{B}$ (at w) as the *sum* of the scores of all $p \in \mathcal{A}$.
- Kenny's Generalizations: (a) allow scores of $-\mathfrak{w}$ and $+\mathfrak{r}$, where $\mathfrak{w} \geq \mathfrak{r} > 0$, and (b) allow S to *suspend on* p [$S(p)$], where all suspensions are given a *neutral* score of *zero*.
- This generalization of our framework leads to an elegant analogue of our central Theorem that $(\mathcal{R})$ entails (WADA).

Theorem. An agent S will avoid (strict) dominance in *total score* ***if*** their belief set $\mathbf{B}$ can be represented as follows:

$(\mathfrak{R})$ There exists a probability function $\Pr(\cdot)$ such that, $\forall p \in \mathcal{A}$:

$$B(p) \textit{ iff } \Pr(p) > \tfrac{\mathfrak{w}}{\mathfrak{r}+\mathfrak{w}},$$

$$D(p) \textit{ iff } \Pr(p) < 1 - \tfrac{\mathfrak{w}}{\mathfrak{r}+\mathfrak{w}},$$

$$S(p) \textit{ iff } \Pr(p) \in \left[1 - \tfrac{\mathfrak{w}}{\mathfrak{r}+\mathfrak{w}}, \tfrac{\mathfrak{w}}{\mathfrak{r}+\mathfrak{w}}\right].$$

- We (along with Rachael Briggs and Fabrizio Cariani) [2] are investigating various applications of this new approach.
- One interesting application is to *judgment aggregation*. *E.g.*,
  - Majority rule aggregations of the judgments of a bunch of agents — each of whom satisfy (PV) — *need not* satisfy (PV).
- **Q**: does majority rule preserve *our* notion of coherence, *viz.*, is (WADA) preserved by MR? **A**: yes (on simple, atomic + truth-functional agendas), but *not on all possible agendas*.
  - There are (not merely atomic + truth-functional) agendas A and sets of judges J ($|A| \geq 5$, $|J| \geq 5$) that (severally) satisfy (WADA), while their majority profile *violates* (WADA).
- *But*, if a set of judges is (severally) *consistent* [*i.e.*, satisfy (PV)], then their majority profile *must* satisfy (WADA).

☞ **Recipe**. Wherever **B**-*consistency* runs into paradox, substitute *coherence* (in *our* sense), and see what happens.

[1] L. Bonjour, *The Coherence Theory of Empirical Knowledge*, *Phil. Studies*, 1975.

[2] R. Briggs, F. Cariani, K. Easwaran, B. Fitelson, *Individual Coherence and Group Coherence*, to appear in *Essays in Collective Epistemology*, J. Lackey (*ed.*), OUP.

[3] J. Broome, *Wide or Narrow Scope*, *Mind*, 2007.

[4] M. Caie, *Rational Probabilistic Incoherence*, to appear in *The Philosophical Review*.

[5] R. Carnap, *Logical Foundations of Probability*, U. of Chicago, 2nd ed., 1962.

[6] D. Christensen, *Putting Logic in its Place*, OUP, 2007.

[7] W. Clifford, *The ethics of belief*, *Contemporary Review*, 1877.

[8] M. Deza and E. Deza, *Encyclopedia of Distances*, Springer, 2009.

[9] C. Duddy and A. Piggins, *A measure of distance between judgment sets*, *Social Choice and Welfare*, 2012.

[10] K. Easwaran, *Dr. Truthlove or: How I Learned to Stop Worrying and Love Bayesian Probability*, manuscript.

[11] K. Easwaran and B. Fitelson, *Accuracy, Coherence, and Evidence*, to appear in *Oxford Studies in Epistemology V*, T. Szabo-Gendler and J. Hawthorne (*eds.*), OUP.

[12] B. Fitelson, *Coherence*, book manuscript, 2014.

[13] B. Fitelson and D. McCarthy, *Toward an Epistemic Foundation for Comparative Confidence*, manuscript, 2014.

[14] R. Foley, *Working Without a Net*, Oxford University Press, 1992.

[15] R. Fumerton, *Metaepistemology and Skepticism*, Rowman & Littlefield, 1995.

[16] A. Hájek, *A Puzzle About Degree of Belief*, manuscript.

[17] ______, *Arguments for — or Against — Probabilism?*, BJPS, 2008.

[18] W. James, *The will to believe*, *The New World*, 1896.

[19] J. Joyce, *Accuracy and Coherence: Prospects for an Alethic Epistemology of Partial Belief*, in F. Huber and C. Schmidt-Petri (*eds.*), *Degrees of Belief*, 2009.

[20] ______, *How Probabilities Reflect Evidence*, *Philosophical Perspectives 19*, 2005.

[21] ______, *A Nonpragmatic Vindication of Probabilism*, *Philosophy of Science*, 1998.

[22] M. Kaplan, *Coming to Terms with our Human Fallibility*, *PPR*, 2013.

[23] P. Klein, *The virtues of inconsistency*, *The Monist*, 1985.

[24] N. Kolodny, *How Does Coherence Matter?*, *Proc. of the Aristotelian Society*, 2007.

[25] N. Kolodny and J. MacFarlane, *Ifs and oughts*, *Journal of Philosophy*, 2010.

[26] D. Miller, *Out of error: Further essays on critical rationalism*, Ashgate, 2006.

[27] D. Nelkin, *The lottery paradox, knowledge, and rationality*, *Phil. Review*, 2000.

[28] J. Pollock, *Nomic probability and the foundations of induction*, OUP, 1990.

[29] F. Ramsey, *Truth and Probability*, 1928.

[30] S. Ryan, *The epistemic virtues of consistency*, *Synthese*, 1996.

[31] M. Titelbaum, *Quitting Certainties*, OUP, 2013.

[32] T. Williamson, *Knowledge and its Limits*, Oxford University Press, 2000.